

Applicant: Deutsches Krebsforschungszentrum  
Our ref: K 2991 - sch / msl

### **Novel regulatory sequences of the MCP-1 gene**

The present invention relates to a nucleic acid molecule comprising (a) a nucleic acid sequence encoding the monocyte-chemoattractant-protein-1 (MCP-1) or a protein having the biological activity of the monocyte-chemoattractant-protein-1 (MCP-1); and (b) a hypersensitive region. The present invention also relates to pharmaceutical compositions for the treatment of diseases associated with dysregulation of MCP-1 expression, e.g. atherosclerosis or cancer. These pharmaceutical compositions comprise either compounds directly or indirectly interacting with the hypersensitive regions or are based on the replacement of the original hypersensitive sequences by mutated versions thereof.

The monocyte-chemoattractant-protein-1 (MCP-1) triggers the infiltration and activation of cells of the monocyte-macrophage lineage, representing known producers of growth-modulatory cytokines. Synthesis of cytokines can be considered an important regulatory loop in the immunological surveillance, since  $\text{TNF-}\alpha$  or  $\text{TGF-}\beta$  can negatively interfere with the expression of high risk human papillomaviruses (HPV) types, the causative agents of cancer of the cervix uteri. Besides disturbance of a functional T-cell surveillance, dys-regulation of chemokine expression may represent another important counter-selection event during the multi-step progression to cervical cancer. Consistent with this anticipation was the recent finding that MCP-1 expression indeed gradually disappears in different stages of premalignant cervical intraepithelial neoplasias, a process, which may explain the severe reduction of intraepithelial macrophages and Langerhans' cells. It is also known that MCP-1 does not only play a critical role during tumor progression but also as regards atherosclerosis. However, due to the fact that so far the knowledge about the complex regulation of MCP-1 gene expression and regulatory regions involved was

09899276 070601

limited, a therapeutically useful modulation of MCP-1 gene expression was not possible.

Thus, the technical problem underlying the present invention is to provide means for searching active substances for the treatment of diseases associated with an abnormal MCP-1 expression, e.g. atherosclerosis and cancer.

The solution to said technical problem is achieved by providing the embodiments characterized in the claims.

It has been, surprisingly, found that the expression of MCP-1 is dependent on a variety of cis-acting sequences which were so far not known. In the experiments leading to the present invention it could be shown that the expression of MCP-1 is closely linked with a non-tumorigenic phenotype in somatic cell hybrids made between the human papilloma virus type 18 (HPV 18) positive cervical carcinoma cell line HeLa and normal human fibroblasts. In contrast, MCP-1 transcription is absent in tumorigenic segregants derived from the same hybrids or in parental HeLa cells. Selectivity of MCP-1 transcription, which is regulated at the level of initiation of transcription, is mainly based on differences in the location and extension of DNase I-hypersensitive regions (DHSR) at both ends of the gene or S1-hypersensitive sites (SHS). DNA sequencing showed that the hypersensitive sites often coincide with binding sites for transcription factors, like an AP-1 site, SP1 site, NF-IL6 site or NF kappa B site. Analyses of AP-1 composition revealed that MCP-1 is only expressed in those cells where *jun*-family members were mainly heterodimerized with the *fos*-related protein *fra-1*. In contrast, in tumorigenic cells the 1:1 ratio between *jun* and *fra-1* is disturbed and the MCP-1 gene is no longer expressed. Hence, alterations in the heterodimerization pattern of AP-1 and its selective accessibility to opened chromatin may represent a novel regulatory pathway in the regulation of chemokines in malignant and non-malignant HPV-positive cells.

In particular the present invention relates to a nucleic acid molecule comprising

(a) a nucleic acid sequence encoding the monocyte-chemoattractant-protein-1

(MCP-1) or a protein having the biological activity of the monocyte-chemoattractant-protein-1 (MCP-1); and

- (b) a 3'-DHSR comprising a nucleic acid molecule which is located 2430 bp to 3019 bp downstream of the transcriptional start site of the MCP-1 gene, or a 3'-DHSR comprising a nucleic acid molecule which is located 1550 bp to 1749 bp downstream of the transcriptional start site of the MCP-1 gene, or a 3'-DHSR comprising a nucleic acid molecule which is located 750 bp to 899 bp downstream of the transcriptional start site of the MCP-1 gene, or a 5'-DHSR comprising a nucleic acid molecule which is located 500 bp to 251 bp upstream of the transcriptional start site of the MCP-1 gene, or a 5'-DHSR comprising a nucleic acid molecule which is located 1300 bp to 1001 bp upstream of the transcriptional start site of the MCP-1 gene, or a 5'-DHSR comprising a nucleic acid molecule which is located 5050 bp to 4751 bp upstream of the transcriptional start site of the MCP-1 gene, or a S1 hypersensitive site comprising a nucleic acid molecule which is located in the 1st intron (+180 to +350) of the MCP-1 gene

As used herein, the term "a nucleic acid molecule encoding the monocyte-chemoattractant-protein-1 (MCP-1) or a protein having the biological activity of the monocyte-chemoattractant-protein-1 (MCP-1)" relates to the published sequence of the MCP-1 gene, e.g. the EMBL database sequence (accession number: Y 18933), the sequence published by Rollins et al., *Mol. Cell. Biol.* 9 (1989), 4687-4695) or a related sequence which still encodes a protein having at least one of the activities of MCP-1.

These related sequences comprise sequences which hybridize to the above published MCP-1 nucleic acid sequences. As used herein, the term "hybridize" has the meaning of hybridization under conventional hybridization conditions, preferably under stringent conditions as described, for example, in Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 2<sup>nd</sup> edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Also contemplated are nucleic acid sequences that hybridize to the MCP-1 nucleic acid sequences at lower stringency hybridization

conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency), salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37°C in a solution comprising 6X SSPE (20 x SSPE = 3M NaCl; 0.2M  $\text{NaH}_2\text{PO}_4$ ; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100  $\mu\text{g/ml}$  salmon sperm blocking DNA, followed by washes at 50°C with 1 x SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5 x SSC). Variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

These related nucleic acid sequences also include sequences that are degenerate as a result of the genetic code or fragments, derivatives and allelic variants of the nucleic acid molecules described above encoding a MCP-1 protein or a protein having MCP-1 activity. "Fragments" are understood to be parts of the nucleic acid sequences that are long enough to encode one of the described proteins. The term "derivative" in this context means that the sequences of these molecules differ from the sequences of the nucleic acid sequences described above at one or several positions but have a high level of homology to these sequences. Homology hereby means a sequence identity of at least 40%, in particular an identity of at least 60%, preferably of more than 80% and particularly preferred of more than 90%. These proteins encoded by the nucleic acid molecules have a sequence identity to the amino acid sequence encoded by the EMBL clone (Y 18933) of at least 80%, preferably of 85% and particularly preferred of more than 90%, 95%, 97% and 99%. The deviations to the above-described nucleic acid sequences may have been produced by deletion, substitution, insertion or recombination.

The nucleic acid sequences that are homologous to the above-described sequences and that represent derivatives of these sequences usually are variations of these

molecules that represent modifications having the same biological function. They can be naturally occurring variations, for example sequences from other organisms, or mutations that can either occur naturally or that have been introduced by specific mutagenesis. Furthermore, the variations can be synthetically produced sequences. The allelic variants can be either naturally occurring variants or synthetically produced variants or variants produced by recombinant DNA processes.

The hypersensitivity regions as identified in the present invention are depicted in Figure 6. One of the preferred hypersensitive regions is the 3'-DHSR comprising the nucleic acid sequences from position + 2430 bp to + 3019 bp as shown in Fig. 6. In a more preferred embodiment of the present invention, the 3'-DHSR comprises the nucleic acid sequence GGAAGGTTGAGTCAAGGATT. In an even more preferred embodiment of the present invention, the 3'-DHSR comprises the nucleic acid molecule TGAGTCA. As used herein, the 3'-DHSR sequences also comprise sequences containing alterations compared to the particular sequences shown in the figures or to the published sequences which do not lead to a loss of the function as cis-acting element, e.g. as regards gene expression.

The S1 hypersensitivity region (SHS) in the 1st intron (0.6 kb) of MCP-1 has been identified by S1 nuclease mapping (c.f. Figs. 9 and 10). This SHS contains an AP-1 site which has been verified by band-shift analysis and TNF induction.

The present invention also relates to a nucleic acid molecule wherein the hypersensitive sites contain mutations resulting in a modified DNase I hypersensitivity and S1 hypersensitivity, respectively, or altered interaction with transcription factors, in particular altered interaction with AP-1. Generally, by means of conventional molecular biological processes it is possible (see, e.g., Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2<sup>nd</sup> edition Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) to introduce different mutations. One possibility is the production of deletion mutants in which nucleic acid molecules are produced by continuous deletions from the 5'- or 3'-terminus of the DNA sequence. Another possibility is the introduction of single-point mutation at positions where a modification

influences the interaction with transcription factors, e.g. AP-1. The identity of the mutated sequence with the original sequence is at least 40%, in particular at least 60%, preferably more than 80% and particularly preferred more than 90%.

For the manipulation in prokaryotic cells by means of genetic engineering the nucleic acid molecules of the invention or parts of these molecules can be introduced into plasmids allowing a mutagenesis or a modification of a sequence by recombination of DNA sequences. By means of conventional methods (cf. Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory Press, NY, USA) bases can be exchanged and natural or synthetic sequences can be added. In order to link the DNA fragments with each other adapters or linkers can be added to the fragments. Furthermore, manipulations can be performed that provide suitable cleavage sites or that remove superfluous DNA or cleavage sites. If insertions, deletions or substitutions are possible, in vitro mutagenesis, primer repair, restriction or ligation can be performed. As analysis method usually sequence analysis, restriction analysis and other biochemical or molecular biological methods are used.

The invention furthermore relates to vectors containing the nucleic acid molecules of the invention. Preferably, they are plasmids, cosmids, viruses, bacteriophages and other vectors usually used in the field of genetic engineering. Vectors suitable for use in the present invention include, but are not limited to T7-based expression vectors for expression in bacteria, the pMSXND expression vector for expression in mammalian cells and baculovirus-derived vectors for expression in insect cells. Preferably, the nucleic acid molecule of the invention is operatively linked to the regulatory elements in the recombinant vector of the invention that guarantee the transcription and synthesis of an RNA in prokaryotic and/or eukaryotic cells that can be translated. The nucleotide sequence to be transcribed can be operably linked to a promotor like a T7, metallothionein I or polyhedrin promotor.

In a further embodiment, the present invention relates to recombinant host cells transiently or stably containing the nucleic acid molecules or vectors of the invention.

03399276-070601

Preferably, these cells are prokaryotic or eukaryotic cells, for example mammalian cells, bacterial cells, insect cells or yeast cells.

The present invention also relates to a pharmaceutical composition comprising a compound which is capable of regulating the expression of the MCP-1 gene by directly or indirectly interacting with the nucleic acid sequence corresponding to the hypersensitivity regions as defined above, e.g. the transcription factor itself (e.g. AP-1) or a protein associated with the transcription factor thereby modulating its activity, or the vector containing a nucleic sequence with a modified hypersensitivity region as described above. Depending on the desired medical use these compounds may be compounds interacting with the hypersensitivity region, thus, increasing MCP-1 expression, but also compounds which inhibit the activity of compounds naturally interacting with the hypersensitivity region, thus leading to a decreased expression of MCP-1. Such a compound comprises, e.g. inhibitors which can be, for instance, structural analogues of the transcription factor or a protein associated therewith that act as antagonist. Such an inhibitor is, e.g., a synthetic organic chemical, a natural fermentation product, a substance extracted from a microorganism, plant or animal, or a peptid. Such an inhibitor can also be an antibody or a nucleic acid molecule. The term "antibody", preferably, relates to antibodies which consist essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations. Monoclonal antibodies are made from an antigen containing fragments of the transcription factor or a protein associated therewith by methods well known to those skilled in the art (see, e.g., Köhler et al., Nature 256 (1975), 495). As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')<sub>2</sub> fragments) which are capable of specifically binding to protein. Fab and F(ab')<sub>2</sub> fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and may have less non-specific tissue binding than an intact antibody. (Wahl et al., J. Nucl. Med. 24:316-325 (1983).) Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

0369275-070601

In a preferred embodiment, the compound comprised by the pharmaceutical composition is (a) a protein capable of interacting with the transcription factor (e.g. AP-1, SP1, NF kappa B) or (b) a nucleic acid sequence encoding said protein. Preferably said nucleic acid sequence is integrated into a vector as described above or below, preferably a vector useful for gene therapy.

In a more preferred embodiment, the compound comprised by the pharmaceutical composition is *jun* (e.g. *junD*, *junB*), *fra-1*, *ATF-2*, *jab-1*, *fra-2* or a mixture thereof. It is apparent from the Examples below that by the administration of a mixture containing a particular ratio of *jun* and *fra-1* expression of MCP-1 can be manipulated. Probably, it will be sufficient to overexpress one of these factors in cells (e.g. tumor cells) to reactivate the MCP-1 gene.

For administration these compounds are preferably combined with suitable pharmaceutical carriers. Examples of suitable pharmaceutical carriers are well known in the art and include phosphate buffered saline solutions, water, emulsions, such as oil/water emulsions, various types of wetting agents, sterile solutions etc.. Such carriers can be formulated by conventional methods and can be administered to the subject at a suitable dose. Administration of the suitable compositions may be effected by different ways, e.g. by intravenous, intraperitoneal, subcutaneous, intramuscular, topical or intradermal administration. The route of administration, of course, depends on the nature of the disease and the kind of compound contained in the pharmaceutical composition. The dosage regimen will be determined by the attending physician and other clinical factors. As is well known in the medical arts, dosages for any one patient depends on many factors, including the patient's size, body surface area, age, sex, the particular compound to be administered, time and route of administration, the kind of the disease, general health and other drugs being administered concurrently.

The delivery of nucleic acid molecules encoding proteins useful for regulating MCP-1 expression or of nucleic acid molecules with mutated hypersensitivity region



sequences can be achieved by direct application or, preferably, by using a recombinant expression vector such as a chimeric virus containing these compounds or a colloidal dispersion system. By using appropriate vectors the mutated sequences can be inserted into the genome and, e.g., replace the original sequences. By delivering these nucleic acids to the desired target, the intracellular expression of MCP-1 and, thus, the level of MCP-1 can, e.g., be decreased resulting in the inhibition of the negative effects of MCP-1, e.g. as regards atherosclerosis. On the other hand, by using this approach an increased MCP-1 expression (or re-expression of MCP-1) can be achieved, e.g. for the treatment or prevention of tumors, e.g. cervical carcinoma.

The above nucleic acid sequences can be administered directly to the target site, e.g., by ballistic delivery, as a colloidal dispersion system or by catheter to a site in artery. The colloidal dispersion systems which can be used for delivery of the above nucleic acids include macromolecule complexes, nanocapsules, microspheres, beads and lipid-based systems including oil-in-water emulsions, (mixed) micelles, liposomes and lipoplexes. The preferred colloidal system is a liposome. The composition of the liposome is usually a combination of phospholipids and steroids, especially cholesterol. The skilled person is in a position to select such liposomes which are suitable for the delivery of the desired nucleic acid molecule. Organ-specific or cell-specific liposomes can be used, e.g. in order to achieve delivery only to the desired tissue. The targeting of liposomes can be carried out by the person skilled in the art by applying commonly known methods. This targeting includes passive targeting (utilizing the natural tendency of the liposomes to distribute to cells of the RES in organs which contain sinusoidal capillaries) or active targeting (for example by coupling the liposome to a specific ligand, e.g., an antibody, a receptor, sugar, glycolipid, protein etc., by well known methods). In the present invention monoclonal antibodies are preferably used to target liposomes to specific tumors via specific cell-surface ligands.

Preferred recombinant vectors useful for gene therapy are viral vectors, e.g. adenovirus, herpes virus, vaccinia, or, more preferably, an RNA virus such as a

retrovirus. Even more preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of such retroviral vectors which can be used in the present invention are: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV) and Rous sarcoma virus (RSV). Most preferably, a non-human primate retroviral vector is employed, such as the gibbon ape leukemia virus (GaLV), providing a broader host range compared to murine vectors. Since recombinant retroviruses are defective, assistance is required in order to produce infectious particles. Such assistance can be provided, e.g., by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR. Suitable helper cell lines are well known to those skilled in the art. Said vectors can additionally contain a gene encoding a selectable marker so that the transduced cells can be identified. Moreover, the retroviral vectors can be modified in such a way that they become target specific. This can be achieved, e.g., by inserting a polynucleotide encoding a sugar, a glycolipid, or a protein, preferably an antibody. Those skilled in the art know additional methods for generating target specific vectors. Further suitable vectors and methods for in vitro- or in vivo-gene therapy are described in the literature and are known to the persons skilled in the art; see, e.g., WO 94/29469 or WO 97/00957.

### Brief description of the drawings

#### Figure 1:

**Panel A:** Selective induction and time course of MCP-1 expression.  $5 \times 10^5$  cells were incubated in the presence of 250 U/ml of  $\text{TNF-}\alpha$  for 5 and 18 hours, respectively. 5  $\mu\text{g}$  RNA of each HeLa cells, non-tumorigenic HeLa-fibroblast hybrids ("444") and their tumorigenic segregants ("CGL3") was separated in a 1% agarose gel. After blotting, the filter was consecutively hybridized with cDNA probes encoding the MCP-1 and the  $\beta$ -actin gene. (-): untreated control cells. The positions of the 18 and 28S rRNAs are indicated. **Panel B:** cell morphology of HeLa-, "444"- and "CGL3" cells either untreated (A, C, E) or after simultaneous incubation with  $\text{TNF-}\alpha$  (500 U/ml) and cycloheximide (10  $\mu\text{g}/\text{ml}$ ) for 4 hours at  $37^\circ\text{C}$ . Phase-contrast microscopy

(Magnification: 100 x). **Panel C:** degradation and re-synthesis of I $\kappa$ B- $\alpha$  after treatment of HeLa-, "444"-, and "CGL3"-cells with 250 U/ml TNF- $\alpha$  for 30 minutes (min) or 5 hours (h). 30  $\mu$ g total cellular protein was separated in a 12% SDS-PAGE gel for Western blot analysis. After electrotransfer, the filter was probed with a polyclonal antibody directed against I $\kappa$ B- $\alpha$ . To confirm equal loading, an identical filter was incubated with a  $\beta$ -actin specific antibody.

#### Figure 2: Nuclear run-on analysis

Ongoing transcription from TNF- $\alpha$  treated (250 U/ml, for 5 hours) and control nuclei was examined by hybridizing equal counts of  $^{32}$ P-UTP nascent RNA to nitrocellulose strips carrying the same amount of heat-denaturated DNA probes (2.5  $\mu$ g of MCP-1, HPV 18,  $\beta$ -actin and p"Bluescript" plasmid DNA). To confirm identical hybridization conditions, 50 ng of total cellular DNA was loaded as an internal reference. (-): untreated cells; (+) cells incubated with 250 U/ml of TNF- $\alpha$ .

#### Figure 3: DNase I mapping analysis of the 5'-regulatory region

"444" cells (panel A) and "CGL3" cells (panel B) were treated for 5 hours with TNF- $\alpha$  (250 U/ml). Isotonic nuclei equivalent to 100  $\mu$ g of total DNA from either untreated or TNF- $\alpha$  treated cells were subsequently digested with increasing amounts of DNase I (1.0; 2.0; 3.0; 4.0; and 5.0 units of DNase I/100  $\mu$ l volume) for 5 minutes at 37°C. After extraction, the DNAs were cleaved with Kpn I, separated on 1% agarose gels and hybridized with a cloned PCR fragment located adjacent to Kpn I (pos. -539 to +52 relative to the transcriptional start site defined by Ueda and coworkers (Ueda et al., J. Immunol. **153** (1994), 2052-2063) for indirect end-labeling. The length of the subbands generated after DNase I/Kpn I digestion is indicated in kb. "Cont 1": nuclei lysed immediately after preparation; "Cont 2": DNA from nuclei incubated in digestion buffer without DNase I. The schematic presentation below shows the location of the major and minor DNase I hypersensitive regions (DHSR) relative to the Kpn I site (pos. +211/+216) in both cell lines (indicated by an arrow). The length of the parental 7.7 kb Kpn I/Kpn I fragment and the position of hybridization probe ("probe") are indicated.

Figure 4: Accessibility of the DHSRs to restriction endonucleases in native chromatin

**Panel A:** isotonic nuclei corresponding to an amount of 100  $\mu$ g total cellular DNA from untreated (-) and TNF- $\alpha$  treated (250 U/ml, for 5 hours) "444"- and "CGL3"-Cells were digested with increasing amounts of Dra I (25 and 50 units for 15 minutes at 37°C). After restriction enzyme digestion, the DNAs were processed as described in Fig. 3. "Cont": DNA from nuclei incubated in digestion buffer for 15 minutes at 37°C without any enzyme. The schematic presentation shows the length and the derivation of the subbands generated after Dra I/Kpn I cleavage (for details, see Fig. 3). **Panel B:** same as panel A, only after treatment of the nuclei with Apa I restriction endonuclease.

Figure 5: DNase I mapping analysis of the 3'-regulatory region

"444" cells (**panel A**) and "CGL3" cells (**panel B**) were treated with TNF- $\alpha$  as described in Fig. 3. Nuclei corresponding to 100  $\mu$ g of total DNA obtained from untreated and cytokine treated cells were digested with increasing amounts of DNase I (1.0; 2.0, 3.0, 4.0, and 5.0 units of DNase I/100  $\mu$ l volume) for 5 minutes at 37°C. After lysis, the purified DNAs were cleaved with Hind III, separated on 1% agarose gels and hybridized with a cloned PCR fragment (pos. -539 to +52 relative to the transcriptional start site) located adjacent to the Hind III site for indirect end-labeling. The length of the subbands generated after DNase I/Hind III digestion is indicated in kb. "Cont 1": nuclei immediately after preparation; "Cont 2": DNA from nuclei incubated in digestion buffer without adding enzyme. The schematic presentation below shows the location (indicated by an arrow) of the major and minor DNase I hypersensitive regions (DHSR) relative to the Hind III site (pos -428/-423, based on the transcriptional start site). The length of the parental 4.7 kb Hind III/Hind III fragment, the position of hybridization probe ("probe"), the three exons ("Exon I-III") and the location of the chemokine specific polyadenylation signal ("poly-A") are indicated.

Figure 6: Nucleotide sequence analysis of the 5'- and 3' regulatory region of the human MCP-1 gene covered by DNase I hypersensitive regions (DHSRs) in native chromatin

Nucleotide numbering was done relative to the transcriptional start site (Ueda et al., 1994, J Immunol. 153, p. 2052-2063). Some DHSRs coincide with a DNA fragment shown to harbor functional binding sites for the transcriptional factors, AP-1 and NF- $\kappa$ B (see nucleotide stretch ranging from pos. -2799 to -300). The positions of additional putative binding sequences for the transcription factors AP-1, NF-1L-6, Sp1 and the restriction endonucleases Apa I and Dra I are indicated. The complete nucleotide sequence (11793 bp) of the whole MCP-1 gene locus is available at the EMBL nucleotide sequence data base (accession number: Y 18933).

Figure 7: Band shift analysis of the 3'AP-1 site within the major DHSR

**Panel A:** 2.5  $\mu$ g "444" cell extract was incubated with a  $^{32}$ P-labeled oligonucleotide from pos. +2587/+2607 (referred as AP-1 (3')), see schematic presentation depicted below). Lane 1: without competition; lane 2: addition of a 100-fold molar excess of unlabeled homologous oligonucleotides ("hom. comp. AP-1 3'"); lane 3: addition of a 100-fold molar excess of an oligonucleotide harbouring a recognition site for the transcription factor NF-1 ("hetero. comp. NF-1"); lane 4: competition with a 100-fold excess of an oligonucleotide containing an AP-1 site of the human collagenase gene ("hom. comp. consensus"). **Panel B:** (lanes 1-4) band-shift analysis using the AP-1 3'-oligonucleotide ("AP-1 3'-end") after incubation with nuclear extracts derived from control (-) and TNF- $\alpha$  treated (+) "444"- and "CGL3" cells. Lanes 5 and 6: band-shift pattern of an AP-1 oligonucleotide from the human collagenase gene ("AP-1 consens.") using either "444" or "CGL3" nuclear extracts. Cells were treated with TNF- $\alpha$  (250 U/ml) for 5 hours (5h). The position of the AP-1 specific DNA-protein complex is indicated. The schematic presentation below shows the location and sequence of the 3'-AP-1 site within the 3'-major DHSR relative to the genome organization of the MCP-1 gene.

Figure 8:

**Panel A:** Supershift-analysis of the AP-1 complex in "444" cells and "CGL3" cells.  $^{32}$ P-

Figure 9:

Figure 10:

Band-shift analysis of the AP-1 site located within the 1. intron of the MCP-1 gene. A: 2.5  $\mu$ g "444" extract was incubated with a  $^{32}$ P-labeled oligonucleotide from position +247/+267 (GATAAGGTGACTCAGAAAAGG, referred as AP-1(I), see schematic presentation depicted below). Lane 1: without competition, lane 2: addition of a 100-fold molar excess of unlabeled homologous oligonucleotides ("homol. comp.(AP-1(I));

B: First four lanes represent band-shift analysis using the AP-1 (I) oligonucleotide (AP-1 (I)) after incubation with nuclear extracts derived from control (-) and TNF- $\alpha$  treated (+) "444"- and "CGL3" cells. Right two lanes represent band-shift pattern of an AP-1 oligonucleotide from the human collagenase gene (AP-1 cons.) using either "444"- or "CGL3" nuclear extracts. Cells were treated with TNF- $\alpha$  (250 U/ml) for 5 h. The position of the AP-1 specific DNA protein complex is indicated. The schematic presentation below shows the location and the sequence of the AP-1 (I) site within the 1st intron relative to the genome organisation of the MCP-1 gene.

The following examples illustrate the invention.

### EXAMPLE 1: GENERAL METHODS

### (A) Cell lines

HPV 18-positive cervical carcinoma cells (HeLa), non-tumorigenic somatic cell hybrids made between HeLa cells and normal human fibroblasts ("444") (Stanbridge, Cancer Survey 3 (1984), 335-350), the corresponding tumorigenic segregants ("CGL3") (Stanbridge, Cancer Survey 3 (1984), 335-350) and the human glioblastoma cell line A172 (Tumorbank DKFZ, Heidelberg) were maintained in Dulbecco's modified Eagle's medium, supplemented with 10% fetal calf serum and 1% penicillin/streptomycin

### (B) Cytokines and reagents

For cytokine treatment, the cells were incubated with TNF- $\alpha$  (specific activity:  $1 \times 10^8$  units/mg) (Pharma Biotechnologie, Hannover, Germany) for different periods of time as described in the figure legends. Cycloheximide (Sigma, St. Louise, USA) was stored as a 10 mg/ml stock solution at  $-20^\circ\text{C}$ .

### (C) RNA analysis

Total cellular RNA was extracted according to the guanidine-thiocyanate procedure (Chomczynski and Sacchi, *Anal. Biochem.* **162** (1987), 156-159). Approximately 5  $\mu$ g of total cellular RNA was separated on 1% agarose gels as described elsewhere (Rösl et al., *J. Virol.* **68** (1994), 2142-2150).

### (D) Quantification of apoptosis

The rate of apoptosis was determined by a cell death detection ELISA<sup>PLUS</sup> kit (Roche) as recommended by the manufacturer. The enrichment factor was calculated by setting the absorbance of the untreated cells as 1.

### (E) Isolation and DNA sequence analysis of the human MCP-1 gene

A chromosome 17 specific cosmid library (Resourcenzentrum Berlin; Zehetner, G. Lehrach, H. (1994), *Nature* **367**, p. 489-491) was screened using the cDNA of MCP-1 as a probe. The sequence of the flanking genomic regions was finally determined using two plasmid clones harboring overlapping inserts of about 7.7 kb (a Kpn I/Kpn I-fragment for the 5'-region) and a 4.7 kb (Hind III/Hind III-fragment for intron/exon and the 3'-region) in size. DNA sequence analysis was done on both strands using chemically synthesized primers and fluorescent sequencing on an automatic sequencer (Model 377, Perkin-Elmer/Applied Biosystems (Foster City, USA) and the dye terminator method according to the manufacturers protocol ("ABI PRISM Big Dye Ready Reaction Terminator Cycle Sequencing Kit", Perkin-Elmer/Applied Biosystems). Putative binding sites for transcription factors were identified by the computer program "Transfac version 3.3" (Wingender et al., *Nucl. Acids. Res.* **25** (1997), 265-268). The complete 11793 bp sequence of the human MCP-1 gene locus encompassing the 5'- and 3'-regulatory regions as well as the introns is available at the EMBL nucleotide sequence data base (accession number: Y18933).

### (F) RNA and DNA probes

HPV 18 represents the unit-length of HPV 18 DNA cloned in pBR322 (Boshart et al., *EMBO J.* **3** (1984), 1151-1157). pHFAI (Gunning et al., *Mol. Cell. Biol.* **3** (1983), 787-795) containing an approximately full-length insert of the fibroblast  $\beta$ -actin gene was



a generous gift from L. Kedes (Medical Center, Palo Alto, USA). The full length cDNA encoding the monocyte-chemoattractant-protein-1 (MCP-1) (Rollins et al., Mol. Cell. Biol. 9 (1989), 4687-4695) was obtained from the American Type Culture Collection (Kockville, MD, USA; ATCC 61364/61365, plasmid clone pGEM-hJE34, HGLM probe ID: p11696). The cDNA of the human type 1 collagenase gene (Offringa et al., Cell 62 (1990), 527-538) was kindly provided by P. Angel (Deutsches Krebsforschungszentrum). The PCR product encompassing parts of the MCP-1 promotor ranges from -539 to +52 (according to Ueda et al., 1994) was amplified and subcloned as described previously (Rösl et al., 1994).

### **(G) SDS-PAGE and Western blotting**

Cellular extracts (Dignam et al., Nucl. Acids Res. 11, (1983), 1475-1489) were separated in sodium dodecyl sulfate (SDS) 12% polyacrylamide gels (PAGE), electrotransferred to PVDF membranes (Immobilon-P, Millipore Corporation Bedford, MA, USA) and probed with polyclonal rabbit  $\text{I}\kappa\text{B}\alpha$  antibodies (epitope corresponding to the amino acids 297-317 mapping at the carboxy terminus, Santa Cruz Biotechnology, California, USA) as described recently (Soto et al., Oncogene 18 (1999), 3187-3198). Equal protein transfer and loading was routinely monitored by incubating identical filters with a monoclonal  $\beta$ -actin specific antibody (ICN Biomedicals, Ohio, USA).

### **(H) Nuclear run-on assays**

Nuclei from exponentially growing cells were prepared as described recently (Maehama et al., Int. J. Cancer 76 (1998), 639-646). Heat-denaturated linearized DNA fragments were fixed onto nitrocellulose filters using the "S+S minifold II slot apparatus" (Schleicher and Schuell, Dassel, Germany) and hybridized with equal amounts of  $^{32}\text{P}$ -labelled RNA. After 48h, the filters were washed with 2 x SSC/0.1% SDS at 68°C.

### **(I) Isolation of nuclei and nuclease digestion conditions for chromatin analysis**

Nuclei from HeLa-, "CGL3"- and "444"-cells were prepared (Rösl et al., Mol. Carcinog. 2 (1989), 72-80) and digested as described in the figure legends. Southern blot

analyses were done using standard protocols.

#### (J) Electrophoretic mobility shift assays

To analyse protein/DNA interactions by electrophoretic mobility shift assays (EMSAs), the following oligonucleotides were used: the AP-1 binding site (5'-GGAAGGTTGAGTCAAGGGATT-3') located within the DNase I hypersensitive site downstream of the MCP-1 gene (pos. +2587 / +2607), an AP-1 consensus sequence, 5'-CGCTTGATGACTCAGCCGGAA-3' (Lee et al., Cell 49 (1987), 741-752), an oligonucleotide containing a recognition site for the nuclear factor 1 (NF-1) derived from the adenovirus origin, 5'-TTTTGGATTGAAGCCAATGATAA-3' (Kenny and Hurwitz, J. Biol. Chem. 263 (1988), 9809-9817). The DNAs were synthesized using a phosphoramidite chemistry (Applied Biosystems synthesizer) and further purified by HPLC. Preparation of nuclear extracts, electrophoretic mobility shift and supershift assays were performed exactly as described (Soto et al., Oncogene 18 (1999), 3187-3198).

#### **EXAMPLE 2: Selective MCP-1 inducibility by TNF- $\alpha$ : demonstration of the integrity of the TNF- $\alpha$ signal transduction pathway in malignant and non-malignant HPV 18-positive cells**

By monitoring the expression of the MCP-1 gene in HPV 18 positive cells differing in their potential to form tumors in nude mice, it was found that MCP-1 was transcribed and TNF- $\alpha$  inducible only in non-malignant hybrids made between HeLa cervical carcinoma cells and normal human fibroblasts (Fig. 1A, designated as "444"). In contrast, examination of either the tumorigenic segregants from the same hybrids (designated as "CGL3") or the parental tumorigenic HeLa cells showed the complete absence of detectable transcripts under the same experimental conditions. The intensity of the MCP-1 mRNA reached a maximum 5 hours after TNF- $\alpha$  addition and declined to almost baseline level 18 hours after TNF- $\alpha$  treatment. The absence of MCP-1 inducibility in the tumorigenic cells cannot be attributed to disturbances in the cytokine signal transduction pathway, since the simultaneous application of TNF- $\alpha$  and a protein synthesis inhibitor (cycloheximide) for 4 hours induced blebbing of the cellular membrane (Fig. 1B), a morphological feature characteristic for TNF- $\alpha$ .

mediated apoptosis. The extend of apoptosis was quantified using a cell death detection ELISA kit which determines the amount of free histones within the cytoplasmic fraction after nuclear envelope breakdown and DNA fragmentation into discrete multiplicities of mononucleosomes (Fig. 1C). The integrity of the TNF- $\alpha$  pathway was further confirmed by the rapid proteolysis (after 30 minutes) and delayed resynthesis (after 5 hours) of the inhibitory component I $\kappa$ B- $\alpha$  (Fig. 1D), leading to the activation of the transcription factor NF- $\kappa$ B.

### **EXAMPLE 3: MCP-1 mRNA expression is regulated at the level of initiation of transcription**

In order to address the question of whether the restricted expression of MCP-1 and its TNF- $\alpha$  inducibility in non-tumorigenic cells was due to a transcriptional block, nuclear run-on experiments were carried out. Hybridization of equal counts of  $^{32}$ P-UTP labeled total RNA to defined amounts of immobilized DNA showed the absence of MCP-1 expression in TNF- $\alpha$  treated tumorigenic "CGL3" cells, while the gene became strongly induced in "444" cells 5 hours after TNF- $\alpha$  addition (Fig. 2). The induction was selective, which was confirmed by a direct comparison with the degree of transcriptional activity of other reference genes such as HPV 18 or  $\beta$ -actin. Furthermore, loading a small amount of total cellular DNA on nitrocellulose strips demonstrated that the total  $^{32}$ P-UTP incorporation of the nascent RNA was the same. These data indicate that TNF- $\alpha$  mediated MCP-1 induction is regulated at the level of initiation of transcription rather than by post-transcriptional mechanisms.

### **EXAMPLE 4: Mapping of DNase I hypersensitive regions (DHSRs) at the 5'- and 3'-end of the MCP-1 gene in native chromatin: evidence for a novel TNF- $\alpha$ inducible DHSR at the 3'-end**

In order to obtain more information on MCP-1 transcriptional regulation, the chromatin structure of the whole MCP-1 locus was examined. As shown in Figure 3A, there is one major DNase I hypersensitive region (DHSR) in "444"-cells, whose maximum DNase I cutting frequency was located approximately 2.4-2.9 kb upstream of the Kpn I site (Pos. +211/+216 relative to the transcriptional start site; see Ueda et al., 1994). In contrast, the major DHSR visible in "CGL3" cells has a smaller extension only

ranging from 2.7-2.9 kb (Fig. 3B). In addition, the 5'-chromatin structure in "444"-cells revealed the presence of two minor DHSRs (see schematic presentation below): a distal and a proximal one, located 5.1 and 1.5 kb, respectively, upstream of the Kpn I site (for DNA sequence analysis, see Fig. 6).

Taking advantage that the restriction endonuclease Dra I is located not only within the minor distal DHSR (pos. -1259/-1254) but also at the 3'-boundary of the major "444"-DHSR, (pos. -2186/-2191, see also Fig. 6), incubation of isolated nuclei with Dra I provides additional evidence for the larger size and higher TNF- $\alpha$  mediated susceptibility of this chromatin stretch in non-malignant cells. As shown in Fig. 4A, digestion of "444" derived nuclei led to a fast and strong appearance of two Dra I-derived subfragments (2.4 and 1.5 kb, respectively), whose intensity were even pronounced after TNF- $\alpha$  application. Quantifying the resulting subbands relative to the intensity of the undigested parental 7.7 kb fragment showed that already at 15 minutes incubation with 25 units of Dra I was sufficient to cleave almost 45% of the chromatin (corresponding to 100  $\mu$ g DNA). In contrast, the same region was much more resistant in untreated "CGL3" control cells (Fig. 4A). Accessibility to Dra I was only weakly elevated by TNF- $\alpha$  (accumulation of the cleaved subbands increased from 6 to 12% relative to the parental fragment), unequivocally demonstrating a different degree of chromatin decondensation in both cell lines. Treating the same nuclei with Apa I (pos. -2692/-2697), a restriction endonuclease which is located directly within the major 5'-DHSRs in both cell lines, no differences in the temporal appearance and the intensity of the generated subbands could be observed (Fig. 4B).

Cytokine treatment led to a strong appearance of a prominent 3'-DHSR, which is located between 2.7-3.4 kb downstream of the Hind III recognition site (pos. -428/-423). Furthermore, several sensitive regions could be noticed (Fig. 5, see schematic overview). It should be stressed that major 3'-DHSR formation was not merely the consequence of MCP-1 transcription per se, because the chromatin of the untreated "444" control cells was resistant against DNase I (Fig. 5, panel A, left), despite ongoing baseline MCP-1 expression was detectable (see Fig. 1, panel A). Furthermore, TNF- $\alpha$  kinetics and subsequent DNase I mapping analyses indicated

that DHSR configuration even slightly preceded MCP-1 induction, suggesting that changes in the nucleosomal organization was a requisite rather than a consequence of TNF- $\alpha$  mediated MCP-1 transcription. Carrying out the same experimental approach for both the tumorigenic and segregants (Fig. 5B) or the parental HeLa cells, no 3'-DHSR could be detected. One can therefore conclude that the differences of MCP-1 transcription and TNF- $\alpha$  inducibility in tumorigenic and non-tumorigenic HPV-positive cells is obviously mediated by a concerted action of at least two different regulatory regions. One is located at the 5'-end, being characterized by their higher complexity in terms of location and distribution of DHSRs; the other is situated approximately 600 bp downstream of the MCP-1 specific polyadenylation signal within a 3'-regulatory domain, whose chromatin structure only became accessible after cytokine treatment.

**EXAMPLE 5: DNA sequencing and band-shift analysis led to the detection of a novel AP-1 site within the 3'-DNase I hypersensitive region**

Since the MCP-1 gene is located at chromosome 17 a chromosome-specific library was screened to isolate the whole genomic locus of the human MCP-1 gene using the corresponding cDNA (Rollins et al., 1989, Mol. Cell. Biol. 9, p. 4687-4695) and a PCR amplified promoter fragment (Rösl et al., 1994, J. Virol. 68, p. 2142-2150) as hybridization probe. DNA sequence analysis essentially confirmed the sequence data of the 3.8 kb upstream region published by Ueda and co-workers (Ueda et al., 1994) (see also Fig. 6). Single nucleotide variations may be dependent on different origins of the clones used for sequencing. Inspecting the nucleotide sequence surrounding the minor 5'-DHSR 1.5 kb upstream of the Kpn I site, a putative NF-IL-6 (C/EBP- $\beta$ ) consensus sequence was found (Fig. 6). Similar sequences could be detected within the minor 2.0 kb subfragment and the major 3'-end DHSR 2.7-3.4 bp subfragments (see Fig. 6). The nucleotide stretch covered by the major DHSR in native "444"-chromatin, where two contiguous NF- $\kappa$ B sites and one AP-1 site are present, could be detected (Fig. 6). Since all three cis-regulatory sequences were shown to be functional in DNA binding assays, they were not further followed up in this experiment. Both the distal 5.1 kb and the proximal 1.5 kb minor 5'-DHSRs coincide with putative consensus sequences for the transcription factor Sp 1.

Scanning the 3'-end for potential regulatory sequences, an additional downstream AP-1 site (5'-TGAGTCA-3'; pos. +2594/+2600) (Fig. 6) could be detected. To prove the functionality of this sequence in terms of DNA binding, electrophoretic band-shift assays were performed. Using the oligonucleotide 5'-GGAAGGTTGAGTCAAGGGATT-3' (pos. +2587/+2607), a single band can be visualized after incubation with a nuclear extract obtained from "444"-cells (Fig. 7A, lane 1). The specificity of AP-1 binding was confirmed in competition experiments by adding a 100-fold molar unlabeled excess of either the homologous oligonucleotide (lane 2; "hom. comp. AP-1 3'", 5'-GGAAGGTTGAGTCAAGGGATT-3') or a oligonucleotide containing the AP-1 site (lane 4; "comp. AP-1 consensus", 5'-CGCTTGATGACTCAGCCGGAA-3') of the collagenase TPA responsive element, leading in both cases to complete disappearance of the 3'-AP-1 retarded band. No competition, however, could be achieved, when a heterologous, nuclear factor 1 ("NF-1") specific oligonucleotide (lane 3; "hetero.comp. NF-1", 5'-TTTTGGATTGAAGCCAATATGATAA-3') was applied. In order to discern whether or nor the extent of AP-1 binding is different between "444"- and "CGL3"-cells and whether the affinity was modulated by TNF- $\alpha$ , the binding properties of different extracts were analysed in the same polyacrylamide gel (Fig. 7B). Although "CGL3"-cells lack MCP-1 expression (see Fig. 1 and 2), all nuclear extracts showed identical binding affinity independently of whether the cells were treated with TNF- $\alpha$  or not prior to cell harvesting (Fig. 7B, lanes 1-4). The absence of enhanced AP-1 binding was not an exceptional feature of the 3'-AP-1 site, because the affinity to the collagenase specific site was also significantly altered (Fig. 7B, lanes 5 and 6).

Since it could previously be demonstrated that the composition of AP-1 differs considerably with respect to the *jun/fos* and the *jun/fra-1* ratio in tumorigenic and non-tumorigenic cells, band-shift assays in combination with specific antibodies raised against AP- 1 family members were carried out. Antibodies which have been used recognized c-Fos (epitope corresponding to amino acids 3-16 mapping at the N-terminus of human c-Fos protein), Fra-1 (epitope corresponding to amino acids 3-22 mapping at the N-terminus of the human Fra-1 protein), c-Jun (epitope corresponding to amino acids 56-69 mapping within the carboxy-terminal domain of the mouse c-Jun

protein; recognizes both the unphosphorylated and phosphorylated form of c-Jun). All antibodies were obtained from Santa Cruz Biotechnology (Santa Cruz, California, USA) as "TransCruz" supershift reagents (c.f. Soto et al. (1999), Oncogene 18, p. 3187-3198.) As illustrated in Fig. 8A, addition of a *c-jun* antibody resulted in a strong super-shift indicating that there exists substantial amounts of *c-jun* homo- and heterodimers in both cell lines (lanes 2 and 6). A completely different situation, however, was obtained when antibodies against *fra-1* were added after the binding reaction. In "444"-cells, most (50%) of *jun*-family members (e. g. c-jun) were mainly heterodimerized with the *fos*-related protein *fra-1* (Fig. 9A, lane 4), while in tumorigenic "CGL3"-cells only approximately 15% of the radioactivity (as determined by quantification of the radioactivity in a "PhosphorImager") was retarded within the super-shift complex (lane 8). Conversely, AP-1 in "CGL3"-cells contain *c-fos* (between 5-7%) (lane 7), which is not detectable in non-malignant "444"-cells (lane 3), even after prolonged exposure to the autoradiography. EMSA supershift pattern was not significantly altered after TNF- $\alpha$  addition. To unequivocally demonstrate a different functionality of the transcription factor AP-1 upon cytokine addition in the experimental cell system, the expression of a marker gene which is known to be tightly regulated by AP-1 was studied. As already shown for MCP-1 (Fig. 1, panel A), treatment with TNF- $\alpha$  resulted in a strong induction of the type I collagenase exclusively in "444"-cells, while the gene was transcriptionally silent in the tumorigenic segregants of the same hybrids ("CGL3") or in the parentat HeLa cells (Fig. 8, panel B).